arXiv:1210.7412v1 [math.PR] 28 Oct 2012

Weak Averaging of Semilinear Stochastic Differential Equations with Almost Periodic Coefficients

Mikhail Kamenskii*, Omar Mellah† and Paul Raynaud de Fitte‡

October 18, 2012

Abstract

An averaging result is proved for stochastic evolution equations with highly oscillating coefficients. This result applies in particular to equations with almost periodic coefficients. The convergence to the solution of the averaged equation is obtained in distribution, as in previous works by Khasminskii and Vrkoč.

Keywords : averaging methods, stochastic evolution equations, almost periodic solutions

1 Introduction

Since the classical work of N.M. Krylov and N.N. Bogolyubov [13] devoted to the analysis, by the method of averaging, of the problem of the dependence on a small parameter $\varepsilon > 0$ of almost periodic solutions of ordinary differential equation containing terms of frequency of order $\frac{1}{\varepsilon}$, several articles and books have appeared, which develop this method for different kinds of differential equations. See the bibliography in the book of V.Sh. Burd [5], where a list of books related to this problem for deterministic differential

*Department of Mathematics, Voronezh State University, Voronezh, Russia. E-Mail: Mikhail@kam.vsu.ru

†Laboratoire Raphaël Salem, UMR CNRS 6085, University of Rouen, France and Department of Mathematics Faculty of Sciences University Mouloud Mammeri of Tizi-Ouzou, Algeria E-Mail: omellah@yahoo.fr

‡Laboratoire Raphaël Salem, UMR CNRS 6085, University of Rouen, France E-Mail: prf@univ-rouen.fr

equations is presented. We note here that the authors of these papers are greatly influenced by the books of N.N. Bogolyubov and A.Yu. Mitropolskii [4] and M.A. Krasnosel′skiĭ, V.Sh. Burd and Yu.S. Kolesov [12].

The method of averaging has been applied of course to stochastic differential equations, but in general it was applied to the initial problem in a finite interval, see for example [11]. Even in this we can see a great difference with the deterministic case. To ensure the strong convergence in a space of stochastic processes, we must assume such convergence of the stochastic term when $\varepsilon \to 0$, which virtually excludes the consideration of high frequency oscillation of this term. R.Z. Khasminskii [11] has shown, in a finite dimensional setting, that it is possible to overcome this problem if one only looks for convergence *in distribution* to the solution of the averaged equation. Later Ivo Vrkoč [17] generalized this result in a Hilbert space setting, for which the initial problem was at this time already well developed (see for example the book of Da Prato and Zabczyck [9]).

During the last 20 years an intensive study of the problem of existence of almost periodic solutions of stochastic differential equations was performed by A. Arnold, C. Tudor, G. Da Prato and later by P.H. Bezandry and T. Diagana [1, 2, 3]. For the first group, an almost periodic solution means that the stochastic process generates an almost periodic measure on the paths space. The second group claims the existence of square mean almost periodic solutions, but square mean almost periodicity seems to be a too strong property for solutions of SDEs, see counterexamples in [14].

In this paper we propose the averaging principle for solutions to a family of semilinear stochastic differential equations in Hilbert space which are almost periodic in distribution. The second member of these equations contains a high frequency term. Under the Bezandry-Diagana conditions, we establish the convergence in distribution of the solutions of these equations to the solution of the averaged equation in the sense of Khasminskii-Vrkoč.

The paper is organized as follows: The next section is devoted to the notations and preliminaries. We then prove in Section 3 that the solutions of the equations we consider are almost periodic in distribution, when their coefficients are almost periodic. In section 4, we prove the fondamental averaging result of this paper.

2 Notations and Preliminaries

In the sequel, $(\mathbb{H}_1, \|.\|_{\mathbb{H}_1})$ and $(\mathbb{H}_2, \|.\|_{\mathbb{H}_2})$ denote separable Hilbert spaces and $L(\mathbb{H}_1, \mathbb{H}_2)$ (or $L(\mathbb{H}_1)$ if $\mathbb{H}_1 = \mathbb{H}_2$) is the space of all bounded linear

operators from $\mathbb{H}_1$ to $\mathbb{H}_2$, whose norm will be denoted by $\|.\|_{L(\mathbb{H}_1,\mathbb{H}_2)}$. If $A \in L(\mathbb{H}_1)$ then A^* denotes its adjoint operator and if A is a nuclear operator,

$$|A|_{\mathcal{N}} = \sup\left\{\sum_i |< Ae_i, f_i >|, \{e_i\}, \{f_i\} \text{ orthonormal bases of } \mathbb{H}_1\right\}$$

is the nuclear norm of A.

2.1 Almost periodic functions

Let $(\mathbb{E}, d)$ be a separable metric space, we denote by $C_b(\mathbb{E})$ the Banach space of continuous and bounded functions $f : \mathbb{E} \to \mathbb{R}$ with $\| f \|_\infty = \sup_{x\in\mathbb{E}} |f(x)|$ and by $\mathcal{P}(\mathbb{E})$ the set of all probability measures onto σ-Borel field of $\mathbb{E}$. For $f \in C_b(\mathbb{E})$ we define

$$\| f \|_L = \sup\Big\{\frac{f(x) - f(y)}{d_{\mathbb{E}}(x, y)} : x \neq y\Big\}$$
$$\| f \|_{\mathrm{BL}} = \max\{\| f \|_\infty, \| f \|_L\}$$

and we define

$$\mathrm{BL}(\mathbb{E}) = \big\{f \in C_b(\mathbb{E}); \| f \|_{\mathrm{BL}} < \infty\big\}.$$

For $\mu, \nu \in \mathcal{P}(\mathbb{E})$ we define

$$d_{\mathrm{BL}}(\mu, \nu) = \sup_{\|f\|_{\mathrm{BL}}\le 1}\Big|\int_{\mathbb{E}} f d(\mu - \nu)\Big|$$

which is a complete metric on $\mathcal{P}(\mathbb{E})$ and generates the narrow (or weak) topology, i.e. the coarsest topology on $\mathcal{P}(\mathbb{E})$ such that the mappings $\mu \mapsto \mu(f)$ are continuous for all bounded continuous $f : \mathbb{E} \to \mathbb{R}$.

Let $(\mathbb{E}_1, d_1)$ and $(\mathbb{E}_2, d_2)$ be separable and complete metric spaces. Let f be a continuous mapping from $\mathbb{R}$ to $\mathbb{E}_2$ (resp. from $\mathbb{R}\times\mathbb{E}_1$ to $\mathbb{E}_2$). Let $\mathcal{K}$ be a set of subsets of $\mathbb{E}_1$. The function f is said to be *almost periodic* (respectively *uniformly with respect to x in elements of* $\mathcal{K}$) if for every $\varepsilon > 0$ (respectively for every $\varepsilon > 0$ and every subset $K \in \mathcal{K}$), there exists a constant $l(\varepsilon, K) > 0$ such that any interval of length $l(\varepsilon, K)$ contains at least a number τ for which

$$\sup_{t\in\mathbb{R}} d_2(f(t + \tau), f(t)) < \varepsilon$$
$$(\text{respectively } \sup_{t\in\mathbb{R}}\sup_{x\in K} d_2(f(t + \tau, x), f(t, x)) < \varepsilon).$$

A characterization of almost periodicity is given in the following result, due to Bochner:

Theorem 2.1 *Let $f : \mathbb{R} \to \mathbb{H}_1$ be continuous. Then the following statements are equivalent*

- *f is almost periodic.*
- *The set of translated functions $\{f(t+.)\}_{t\in\mathbb{R}}$ is relatively compact in $\mathcal{C}(\mathbb{R};\mathbb{E}_2)$ with respect to the uniform norm.*
- *f satisfies* Bochner's double sequence criterion*, that is, for every pair of sequences $\{\alpha_n'\} \subset \mathbb{R}$ and $\{\beta_n'\} \subset \mathbb{R}$, there are subsequences $(\alpha_n) \subset (\alpha_n')$ and $(\beta_n) \subset (\beta_n')$ respectively with same indexes such that, for every $t \in \mathbb{R}$, the limits*

  $$\lim_{n\to\infty}\lim_{m\to\infty} f(t+\alpha_n+\beta_m) \text{ and } \lim_{n\to\infty} f(t+\alpha_n+\beta_n), \tag{1}$$

  exist and are equal.

Remark 2.2

(i) A striking property of Bochner's double sequence criterion is that the limits in (1) exist in any of the three modes of convergences: pointwise, uniform on compact intervals and uniform on $\mathbb{R}$ (with respect to $d_{\mathbb{E}}$). This criterion has thus the avantage that it allows to establish uniform convergence by checking pointwise convergence.

(ii) The previous result holds for the metric spaces $(\mathcal{P}(\mathbb{E}), d_{\mathrm{BL}})$ and $(\mathcal{P}(\mathcal{C}(\mathbb{R},\mathbb{E})), d_{\mathrm{BL}})$

2.2 Almost periodic stochastic processes

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. Let $X : \mathbb{R}\times\Omega \to \mathbb{H}_2$ be a stochastic process. We denote by $\mu(t)$ the distribution of the random variable $X(t)$. Following Tudor's terminology [16], we say that X has *almost periodic one-dimensional distributions* if the mapping $t \mapsto \mu(t)$ from $\mathbb{R}$ to $(\mathcal{P}(\mathbb{H}_2), d_{\mathrm{BL}})$ is almost periodic.

If X has continuous trajectories, we say that X is *almost periodic in distribution* if the mapping $t \mapsto \text{law}(X(t+.)$ from $\mathbb{R}$ to $\mathcal{P}(C(\mathbb{R};\mathbb{H}_2))$ is almost periodic, where $C(\mathbb{R};\mathbb{H}_2)$ is endowed with the uniform convergence on compact intervals and $\mathcal{P}(C(\mathbb{R};\mathbb{H}_2))$ is endowed with the distance d_{BL}.

Let $L^2(\mathrm{P},\mathbb{H}_2)$ be the space of $\mathbb{H}_2$-valued random variables with a finite quadratic-mean. We say that a stochastic process $X : \mathbb{R} \to L^2(\mathrm{P},\mathbb{H}_2)$ is *square-mean continuous* if, for every $s \in \mathbb{R}$,

$$\lim_{t\to s} E\|X(t)-X(s)\|^2_{\mathbb{H}_2} = 0.$$

We denote by $\mathrm{CUB}\big(\mathbb{R}, L^2(\mathrm{P}, \mathbb{H}_2)\big)$ the Banach space of square-mean continuous and uniformly bounded stochastic processes, endowed with the norm

$$\|X\|_\infty^2 = \sup_{t\in\mathbb{R}} (E\|X(t)\|_{\mathbb{H}_2}^2).$$

A square-mean continuous stochastic process $X : \mathbb{R} \to L^2(\mathrm{P}, \mathbb{H}_2)$ is said to be *square-mean almost periodic* if, for each $\varepsilon > 0$, there exists $l(\varepsilon) > 0$ such that any interval of length $l(\varepsilon)$ contains at least a number τ for which

$$\sup_{t\in\mathbb{R}} E\|X(t+\tau) - X(t)\|_{\mathbb{H}_2}^2 < \varepsilon.$$

The next theorem is interesting for itself, but we shall not use it in the sequel.

Theorem 2.3 *Let $F : \mathbb{R} \times \mathbb{H}_2 \to \mathbb{H}_2$ be an almost periodic function uniformly with respect to x in compact subsets of $\mathbb{H}_2$ such that*

$$\|F(t,x)\|_{\mathbb{H}_2} \le C_1(1 + \|x\|_{\mathbb{H}_2}) \text{ and } \|F(t,x) - F(t,y)\|_{\mathbb{H}_2} \le C_2\|x - y\|_{\mathbb{H}_2}.$$

Then the function

$$\widetilde{F} : \mathbb{R} \times L^2(\mathrm{P}, \mathbb{H}_2) \to L^2(\mathrm{P}, \mathbb{H}_2)$$

(where $\widetilde{F}(t, Y)(\omega) = F(t, Y(\omega))$ for every $\omega \in \Omega$) is square-mean almost periodic uniformly with respect to Y in compact subsets of $L^2(\mathrm{P}, \mathbb{H}_2)$.

Proof Let us prove that for each $Y \in L^2(\mathrm{P}, \mathbb{H}_2)$ the process $\widetilde{F}_Y : \mathbb{R} \to L^2(\mathrm{P}, \mathbb{H}_2)$, $t \mapsto \widetilde{F}(t, Y)$ is almost periodic.

For every $\delta > 0$, there exists a compact subset S of $\mathbb{H}_2$ such that

$$\mathrm{P}\{Y \notin S\} \le \delta.$$

Let $\varepsilon > 0$, then there exist $\delta > 0$ and a compact subset S of $\mathbb{H}_2$ such that $\mathrm{P}\{Y \notin S\} \le \delta$ and

$$\int_{\{Y\notin S\}} \big(1 + \|Y\|_{\mathbb{H}_2}^2\big)\, d\,\mathrm{P} < \frac{\varepsilon}{4C_1}.$$

Since F is almost periodic uniformly with respect to x in the compact subset S, there exists a constant $l(\varepsilon, S) > 0$ such that any interval of length $l(\varepsilon, S)$ contains at least a number τ for which

$$\sup_t \|F(t+\tau, x) - F(t,x)\|_{\mathbb{H}_2} < \frac{\sqrt{\varepsilon}}{\sqrt{2}} \text{ for all } x \in S.$$

We have

$$\begin{aligned}E(\|F(t+\tau,Y)-F(t,Y)\|_{\mathbb{H}_2}^2) &= \int_{\{Y\in S\}} \|F(t+\tau,Y)-F(t,Y)\|_{\mathbb{H}_2}^2 d\,\mathrm{P} \\ &\quad + \int_{\{Y\notin S\}} \|F(t+\tau,Y)-F(t,Y)\|_{\mathbb{H}_2}^2 d\,\mathrm{P} \\ &< \frac{\varepsilon}{2} + 2C_1 \int_{\{Y\notin S\}} \left(1+\|Y\|_{\mathbb{H}_2}^2\right) d\,\mathrm{P} \\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Therefore the process $\widetilde{F}_Y$ is almost periodic. Since $\widetilde{F}$ is Lipschitz, it is almost periodic uniformly with respect to Y in compact subsets of $L^2(\mathrm{P},\mathbb{H}_2)$ (see [10, Theorem 2.10 page 25]). □

Proposition 2.4 *([7, 10]) For any almost periodic function $F:\mathbb{R}\times\mathbb{H}_2\to\mathbb{H}_2$, there exists a continuous function $F_0:\mathbb{H}_2\to\mathbb{H}_2$ such that for each $x\in\mathbb{H}_2$ the mean value*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T F(t,x)dt = F_0(x). \tag{2}$$

Furthermore, if $F(t,x)$ is Lipschitz in $x\in\mathbb{H}_2$ uniformly with respect to $t\in I$, the mapping F_0 is Lipschitz too.

Let $Q\in L(\mathbb{H}_1)$ be a linear operator. Then Q is a bijection from $\mathrm{range}(Q)=Q(\mathbb{H}_1)$ to $(\ker Q)^\perp$. We denote by Q^{-1} the *pseudo-inverse* of Q (see [15, Appendix C] or [9, Appendix B.2]), that is, the inverse of the mapping $(\ker Q)^\perp\to\mathrm{range}(Q)$, $x\mapsto Q(x)$. Note that $\mathrm{range}(Q)$ is a Hilbert space for the scalar product $\langle x,y\rangle_{\mathrm{range}(Q)} = \langle Q^{-1}(x),Q^{-1}(y)\rangle$.

Proposition 2.5 *Let $G:\mathbb{R}\times\mathbb{H}_2\to L(\mathbb{H}_1,\mathbb{H}_2)$ be an almost periodic function and let $Q\in L(\mathbb{H}_1)$ be a self-adjoint nonnegative operator. Let $\mathbb{H}_0=\mathrm{range}(Q^{1/2})$, endowed with $\langle x,y\rangle_{\mathrm{range}(Q^{1/2})} = \langle Q^{-1/2}(x),Q^{-1/2}(y)\rangle$. There exists a continuous function $G_0:\mathbb{H}_2\to L(\mathbb{H}_0,\mathbb{H}_2)$ such that, for every $x\in\mathbb{H}_1$,*

$$\lim_{t\to\infty}\Big|\frac{1}{t}\int_0^t G(s,x)QG^*(s,x)ds - G_0(x)QG_0^*(x)\Big|_{\mathcal{N}} = 0, \tag{3}$$

where $G^*(s,x)=(G(s,x))^*$ and $G_0^*(x)=(G_0(x))^*$.

Proof Observe first that $G_0(x)QG_0^*(x) = (G_0(x)Q^{1/2})(G_0(x)Q^{1/2})^*$, thus $G_0(x)$ does not need to be defined on the whole space $\mathbb{H}_1$, it is sufficient that it be defined on $\mathbb{H}_0$.

Since G is almost periodic, the function $H(s,x) = G(s,x)QG^*(s,x)$ is almost periodic too, with positive self-adjoint nuclear values in $L(\mathbb{H}_2)$. Thus there exists a mapping $H_0 : \mathbb{H}_2 \to L(\mathbb{H}_2)$ such that, for every $x \in \mathbb{H}_2$,

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t G(s,x)QG^*(s,x)\,ds = H_0(x).$$

By e.g. [10, Theorem 3.1], H_0 is continuous. Thus the mapping

$$H_0^{1/2} : \begin{cases} \mathbb{H}_2 & \to \quad L(\mathbb{H}_2) \\ x & \mapsto \quad (H_0(x))^{1/2} \end{cases}$$

is continuous with positive self-adjoint values.

Let $G_0(x) = H_0^{1/2}(x)Q^{-1/2} : \mathbb{H}_0 \to \mathbb{H}_2$. We then have, for every $x \in \mathbb{H}_2$,

$$H_0(x) = G_0(x)Q(G_0(x))^*$$

and G_0 is continuous, which proves (3). $\square$

3 Solutions almost periodic in distribution

We consider the semilinear stochastic differential equation,

$$dX_t = AX_t dt + F(t,X_t)dt + G(t,X_t)dW(t), t \in \mathbb{R} \tag{4}$$

Where $A : \mathrm{Dom}(A) \subset \mathbb{H}_2 \to \mathbb{H}_2$ is a densely defined closed (possibly unbounded) linear operator, $F : \mathbb{R}\times\mathbb{H}_2 \to \mathbb{H}_2$, and $G : \mathbb{R}\times\mathbb{H}_2 \to L(\mathbb{H}_1,\mathbb{H}_2)$ are continuous functions. In what follows we assume that:

(i) $W(t)$ is an $\mathbb{H}_1$-valued Wiener process with nuclear covariance operator Q (we denote by $\operatorname{tr} Q$ the trace of Q), defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}}, \mathrm{P})$.

(ii) $A : \mathrm{Dom}(A) \to \mathbb{H}_2$ is the infinitesimal generator of a C_0-semigroup $(S(t))_{t\geq 0}$ such that there exists a constant $\delta > 0$ with

$$\|S(t)\|_{L(\mathbb{H}_2)} \leq e^{-\delta t}, t \geq 0.$$

(iii) There exists a constant K such that the mappings $F : \mathbb{R} \times \mathbb{H}_2 \to \mathbb{H}_2$ and $G : \mathbb{R} \times \mathbb{H}_2 \to L(\mathbb{H}_1, \mathbb{H}_2)$ satisfy

$$\|F(t,x)\|_{\mathbb{H}_2} + \|G(t,x)\|_{L(\mathbb{H}_1,\mathbb{H}_2)} \leq K(1 + \|x\|_{\mathbb{H}_2})$$

(iv) The functions F and G are Lipschitz, more precisely there exists a constant K such that

$$\|F(t,x) - F(t,y)\|_{\mathbb{H}_2} + \|G(t,x) - G(t,y)\|_{L(\mathbb{H}_1,\mathbb{H}_2)} \leq K\|x - y\|_{\mathbb{H}_2}$$

for all $t \in \mathbb{R}$ and $x, y \in \mathbb{H}_2$.

(v) The mappings F and G are almost periodic in $t \in \mathbb{R}$ uniformly with respect to x in bounded subsets of $\mathbb{H}_2$.

The assumptions in the following theorem are contained in those of Bezandry and Diagana [1, 2]. The result is similar to [8, Theorem 4.3], with different hypothesis and a different proof.

Theorem 3.1 *Let the assumptions (i) - (v) be fulfiled and the constant* $\theta = \dfrac{K^2}{\delta^2}\Big(\dfrac{1}{2} + \operatorname{tr} Q\Big) < 1$. *Then there exists a unique mild solution X to (4) in* $\mathrm{CUB}\big(\mathbb{R}, L^2(\mathrm{P}, \mathbb{H}_2)\big)$. *Furthermore, X has a.e. continuous trajectories, is almost periodic in distribution, and $X(t)$ can be explicitly expressed as follows, for each $t \in \mathbb{R}$:*

$$X(t) = \int_{-\infty}^{t} S(t-s)F\big(s, X(s)\big)ds + \int_{-\infty}^{t} S(t-s)G\big((s), X(s)\big)dW(s).$$

To prove Theorem 3.1, we will use the following result, which is given in a more general form in ([8]:

Proposition 3.2 *([8, Proposition 3.1-(c)]) Let $\tau \in \mathbb{R}$. Let $(\xi_n)_{0 \leq n \leq \infty}$ be a sequence of square integrable $\mathbb{H}_2$-valued random variables. Let $(F_n)_{0 \leq n \leq \infty}$ and $(G_n)_{0 \leq n \leq \infty}$ be sequences of mappings from $\mathbb{R} \times \mathbb{H}_2$ to $\mathbb{H}_2$ and $L(\mathbb{H}_1, \mathbb{H}_2)$ respectively, satisfying* (iii) *and* (iv) *(replacing F and G by F_n and G_n respectively, and the constant K being independent of n). For each n, let X_n denote the solution of*

$$\begin{aligned} X_n(t) &= S(t-\tau)\xi_n \\ &\quad + \int_{\tau}^{t} S(t-s)F_n\big(s, X_n(s)\big)ds + \int_{\tau}^{t} S(t-s)G_n\big((s), X_n(s)\big)dW(s). \end{aligned}$$

Assume that, for every $(t,x) \in \mathbb{R} \times \mathbb{H}_2$*,*

$$\lim_{n\to\infty} F_n(t,x) = F_\infty(t,x), \ \lim_{n\to\infty} G_n(t,x) = G_\infty(t,x),$$
$$\lim_{n\to\infty} d_{\mathrm{BL}}(\mathrm{law}\,(\xi_n, W)\,, \mathrm{law}\,(\xi_\infty, W)) = 0,$$

(the last equality takes place in $\mathcal{P}\left(\mathbb{H}_2 \times C(\mathbb{R}, \mathbb{H}_1)\right)$*). Then we have in* $\mathcal{C}([\tau, T]; \mathbb{H}_2)$*, for any* $T > \tau$*,*

$$\lim_{n\to\infty} d_{\mathrm{BL}}(\mathrm{law}\,(X_n)\,, \mathrm{law}\,(X_\infty)) = 0.$$

Proof of Theorem 3.1 Note that

$$X(t) = \int_{-\infty}^{t} S(t-s)F\big(s, X(s)\big)ds + \int_{-\infty}^{t} S(t-s)G\big((s), X(s)\big)dW(s)$$

satisfies

$$X(t) = S(t-s)X(s) + \int_{s}^{t} S(t-s)F\big(s, X(s)\big)ds + \int_{s}^{t} S(t-s)G\big((s), X(s)\big)dW(s)$$

for all $t \geq s$ for each $s \in \mathbb{R}$, and hence X is a mild solution to (4).

We introduce an operator L by

$$LX(t) = \int_{-\infty}^{t} S(t-s)F\big(s, X(s)\big)ds + \int_{-\infty}^{t} S(t-s)G\big((s), X(s)\big)dW(s).$$

It can be see easily that the operator L maps $\mathrm{CUB}\big(\mathbb{R}, L^2(\mathrm{P}, \mathbb{H}_1)\big)$ into itself.

First step. Let us show that L has a unique fixed point.

$$\begin{aligned}
&\|(LX)(t) - (LY)(t)\|_{\mathbb{H}_2} \\
&\qquad\qquad = \|\int_{-\infty}^{t} S(t-s)[F(s, X(s)) - F(s, Y(s))]ds \\
&\qquad\qquad + \int_{-\infty}^{t} S(t-s)[G(s, X(s)) - G(s, Y(s))]dW(s)\|_{\mathbb{H}_2} \\
&\qquad\qquad \leq \int_{-\infty}^{t} e^{-\delta(t-s)}\|F(s, X(s)) - F(s, Y(s))\|_{\mathbb{H}_2}ds \\
&\qquad\qquad + \|\int_{-\infty}^{t} S(t-s)[G(s, X(s)) - G(s, Y(s))]dW(s)\|_{\mathbb{H}_2}.
\end{aligned}$$

Using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ we obtain

$$
\begin{aligned}
&E\|(LX)(t) - (LY)(t)\|_{\mathbb{H}_2}^2 \\
&\qquad\leq 2E\Big(\int_{-\infty}^{t} e^{-\delta(t-s)} \|F(s,X(s)) - F(s,Y(s))\|_{\mathbb{H}_2} ds\Big)^2 \\
&\qquad\quad + 2E\Big(\|\int_{-\infty}^{t} S(t-s)[G(s,X(s)) - G(s,Y(s))]dW(s)\|_{\mathbb{H}_2}\Big)^2 \\
&\qquad= I_1 + I_2.
\end{aligned}
$$

We have

$$
\begin{aligned}
I_1 &\leq 2\int_{-\infty}^{t} e^{-\delta(t-s)} ds \int_{-\infty}^{t} e^{-\delta(t-s)} E\|F(s,X(s)) - F(s,Y(s))\|_{\mathbb{H}_2}^2 ds \\
&\leq 2K^2 \int_{-\infty}^{t} e^{-\delta(t-s)} ds \int_{-\infty}^{t} e^{-\delta(t-s)} E\|X(s)) - Y(s))\|_{\mathbb{H}_2}^2 ds \\
&\leq 2K^2 \big(\int_{-\infty}^{t} e^{-\delta(t-s)} ds\big)^2 \sup_{s\in\mathbb{R}} E\|X(s)) - Y(s))\|_{\mathbb{H}_2}^2 \\
&\leq \frac{K^2}{2\delta^2} \sup_{s\in\mathbb{R}} E\|X(s)) - Y(s))\|_{\mathbb{H}_2}^2.
\end{aligned}
$$

For I_2, using the isometry identity we obtain

$$
\begin{aligned}
I_2 &\leq 2\operatorname{tr} Q \int_{-\infty}^{t} e^{-2\delta(t-s)} E\|G(s,X(s)) - G(s,Y(s))\|_{L(\mathbb{H}_1,\mathbb{H}_2)}^2 ds \\
&\leq 2\operatorname{tr} Q K^2 \int_{-\infty}^{t} e^{-2\delta(t-s)} E\|X(s) - Y(s)\|_{\mathbb{H}_2)}^2 ds \\
&\leq 2K^2 \operatorname{tr} Q \big(\int_{-\infty}^{t} e^{-2\delta(t-s)} ds\big) \sup_{s\in\mathbb{R}} E\|X(s) - Y(s)\|_{\mathbb{H}_2}^2 \\
&\leq \frac{K^2 \operatorname{tr} Q}{\delta} \sup_{s\in\mathbb{R}} E\|X(s) - Y(s)\|_{\mathbb{H}_2}^2.
\end{aligned}
$$

Thus

$$
E\|(LX)(t) - (LY)(t)\|_{\mathbb{H}_2}^2 \leq I_1 + I_2 \leq \theta \sup_{s\in\mathbb{R}} E\|X(s) - Y(s)\|_{\mathbb{H}_2}^2.
$$

Consequently, as $\theta < 1$, we deduce that L is a contraction operator, hence there exists a unique mild solution to (4) in $\mathrm{CUB}\big(\mathbb{R}, L^2(\mathrm{P}, \mathbb{H}_1)\big)$.

Furthermore, by [9, Theorem 7.4], almost all trajectories of this solution are continuous.

Second step. Let show that X is almost periodic in distribution. We use Bochner's double sequences criterion. Let $(\alpha_n^{'})$ and $(\beta_n^{'})$ be two sequences in $\mathbb{R}$. We show that there are subsequences $(\alpha_n) \subset (\alpha_n^{'})$ and $(\beta_n) \subset (\beta_n^{'})$ with same indexes such that, for every $t \in \mathbb{R}$, the limits

$$\lim_{n\to\infty} \lim_{m\to\infty} \mu(t + \alpha_n + \beta_m) \text{ and } \lim_{n\to\infty} \mu(t + \alpha_n + \beta_n), \tag{5}$$

exist and are equal, where $\mu(t) := \text{law}\,(X)\,(t)$ is the law or distribution of $X(t)$.

Since F and G are almost periodic, there are subsequences $(\alpha_n) \subset (\alpha_n^{'})$ and $(\beta_n) \subset (\beta_n^{'})$ with same indexes such that

$$\lim_{n\to\infty} \lim_{m\to\infty} F(t + \alpha_n + \beta_m, x) = \lim_{n\to\infty} F(t + \alpha_n + \beta_n, x) =: F_0(t, x) \tag{6}$$

and

$$\lim_{n\to\infty} \lim_{m\to\infty} G(t + \alpha_n + \beta_m, x) = \lim_{n\to\infty} G(t + \alpha_n + \beta_n, x) =: G_0(t, x). \tag{7}$$

These limits exist uniformly with respect to $t \in \mathbb{R}$ and x in bounded subsets of $\mathbb{H}_2$.

Set now $(\gamma_n) = (\alpha_n + \beta_n)$. For each fixed integer n, we consider

$$X^n(t) = \int_{-\infty}^{t} S(t-s)F(s+\gamma_n, X^n(s))ds + \int_{-\infty}^{t} S(t-s)G(s+\gamma_n, X^n(s))dW(s)$$

the mild solution of

$$dX^n(t) = AX^n(t)dt + F(t + \gamma_n, X^n(t))dt + G(t + \gamma_n, X^n(t))dW(t) \tag{8}$$

and

$$X^0(t) = \int_{-\infty}^{t} S(t - s)F_0(s, X^0(s))ds + \int_{-\infty}^{t} S(t - s)G_0(s, X^0(s))dW(s)$$

the mild solution of

$$dX^0(t) = AX^0(t)dt + F_0(t, X^0(t))dt + G_0(t, X^0(t))dW(t). \tag{9}$$

Make the change of variable $\sigma - \gamma_n = s$, the process

$$\begin{aligned} X(t + \gamma_n) = &\int_{-\infty}^{t+\gamma_n} S(t + \gamma_n - s)F(s, X(s))ds \\ &+ \int_{-\infty}^{t+\gamma_n} S(t + \gamma_n - s)G(s, X(s))dW(s) \end{aligned}$$

becomes

$$X(t+\gamma_n) = \int_{-\infty}^{t} S(t-s)F(s+\gamma_n, X(s+\gamma_n))ds + \int_{-\infty}^{t} S(t-s)G(s+\gamma_n, X(s+\gamma_n))d\tilde{W}_n(s),$$

where $\tilde{W}_n(s) = W(s+\gamma_n) - W(\gamma_n)$ is a Brownian motion with the same distribution as $W(s)$. Thus the process $X(t+\gamma_n)$ has the same distribution as $X^n(t)$.

Let us show that $X^n(t)$ converges in mean to $X^0(t)$ for each fixed $t \in \mathbb{R}$. Using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ we obtain

$$\begin{aligned}
&E\|X^n(t) - X^0(t)\|^2 \\
&\quad = E\|\int_{-\infty}^{t} S(t-s)[F(s+\gamma_n, X^n(s)) - F_0(s, X^0(s))]ds \\
&\qquad + \int_{-\infty}^{t} S(t-s)[G(s+\gamma_n, X^n(s)) - G_0(s, X^0(s))]dW(s)\|^2 \\
&\quad \leq 2E\|\int_{-\infty}^{t} S(t-s)[F(s+\gamma_n, X^n(s)) - F_0(s, X^0(s))]ds\|^2 \\
&\qquad + 2E\int_{-\infty}^{t} S(t-s)[G(s+\gamma_n, X^n(s)) - G_0(s, X^0(s))]dW(s)\|^2 \\
&\quad \leq 4E\|\int_{-\infty}^{t} S(t-s)[F(s+\gamma_n, X^n(s)) - F(s+\gamma_n, X^0(s))]ds\|^2 \\
&\qquad + 4E\|\int_{-\infty}^{t} S(t-s)[F(s+\gamma_n, X^0(s)) - F_0(s, X^0(s))]ds\|^2 \\
&\qquad + 4E\|\int_{-\infty}^{t} S(t-s)[G(s+\gamma_n, X^n(s)) - G(s+\gamma_n, X^0(s))]dW(s)\|^2 \\
&\qquad + 4E\|\int_{-\infty}^{t} S(t-s)[G(s+\gamma_n, X^0(s)) - G_0(s, X^0(s))]dW(s)\|^2 \\
&\quad \leq I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Now, using (ii), (iv) and the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}
I_1 &= 4E\|\int_{-\infty}^{t} S(t-s)[F(s+\gamma_n, X^n(s)) - F(s+\gamma_n, X^0(s))]ds\|^2 \\
&\le 4E\big(\int_{-\infty}^{t} \|S(t-s)\|\|F(s+\gamma_n, X^n(s)) - F(s+\gamma_n, X^0(s))\|ds\big)^2 \\
&\le 4E\big(\int_{-\infty}^{t} e^{-\delta(t-s)}\|F(s+\gamma_n, X^n(s)) - F(s+\gamma_n, X^0(s))\|ds\big)^2 \\
&\le 4\big(\int_{-\infty}^{t} e^{-2\delta(t-s)}ds\big)\big(\int_{-\infty}^{t} E\|F(s+\gamma_n, X^n(s)) - F(s+\alpha_n, X^0(s))\|^2 ds\big) \\
&\le \frac{2K^2}{\delta}\int_{-\infty}^{t} E\|X^n(s) - X^0(s)\|^2 ds.
\end{aligned}$$

Then we obtain

$$\begin{aligned}
I_2 &= 4E\|\int_{-\infty}^{t} S(t-s)[F(s+\gamma_n, X^0(s)) - F_0(s, X^0(s))]ds\|^2 \\
&\le 4E\big(\int_{-\infty}^{t} e^{-\delta(t-s)}\|F(s+\gamma_n, X^0(s)) - F_0(s, X^0(s))\|ds\big)^2 \\
&\le 4E\big(\int_{-\infty}^{t} e^{-\delta(t-s)}ds\big)\big(\int_{-\infty}^{t} e^{-\delta(t-s)}\|F(s+\gamma_n, X^0(s)) - F_0(s, X^0(s))\|^2 ds\big) \\
&\le 4\big(\int_{-\infty}^{t} e^{-\delta(t-s)}ds\big)^2 \sup_s E\|F(s+\gamma_n, X^0(s)) - F_0(s, X^0(s))\|^2 \\
&\le \frac{4}{\delta^2}\sup_s E\|F(s+\gamma_n, X^0(s)) - F_0(s, X^0(s))\|^2,
\end{aligned}$$

which converges to 0 as $n \to \infty$ because $\sup_{t\in\mathbb{R}} E\|X^0(t)\|^2 < \infty$ which implies that $\{X^0(t)\}_t$ is tight relatively to bounded sets.

Applying Itô's isometry, we get

$$\begin{aligned}
I_3 &= 4E\|\int_{-\infty}^{t} S(t-s)[G(s+\gamma_n, X^n(s)) - G(s+\gamma_n, X^0(s))]dW(s)\|^2 \\
&\le 4\,\mathrm{tr}\,QE\int_{-\infty}^{t} \|S(t-s)\|^2\|G(s+\gamma_n, X^n(s)) - G(s+\gamma_n, X^0(s))\|^2 ds \\
&\le 4\,\mathrm{tr}\,Q\int_{-\infty}^{t} E\|G(s+\gamma_n, X^n(s)) - G(s+\gamma_n, X^0(s))\|^2 ds \\
&\le 4^2K\,\mathrm{tr}\,Q\int_{-\infty}^{t} E\|X^n(s) - X^0(s)\|^2 ds.
\end{aligned}$$

and

$$
\begin{aligned}
I_4 &= 4E\|\int_{-\infty}^{t} S(t-s)[G(s+\gamma_n, X^0(s)) - G_0(s, X^0(s))]dW(s)\|^2 \\
&\le 4\operatorname{tr} QE\big(\int_{-\infty}^{t} \|S(t-s)\|^2 \|G(s+\gamma_n, X^0(s)) - G_0(s, X^0(s))\|^2 ds\big) \\
&\le 4\operatorname{tr} Q\big(\int_{-\infty}^{t} e^{-2\delta(t-s)} ds\big) \sup_{s\in\mathbb{R}} E\|G(s+\gamma_n, X^0(s)) - G_0(s, X^0(s))\|^2 \\
&\le \frac{2\operatorname{tr} Q}{\delta} \sup_{s\in\mathbb{R}} E\|G(s+\gamma_n, X^0(s)) - G_0(s, X^0(s))\|^2.
\end{aligned}
$$

For the same reason as for I_2, the right hand term goes to 0 as $n\to\infty$. Thus, applying Gronwall's inequality, we obtain

$$\lim_{n\to\infty} E\|X^n(t) - X^0(t)\|^2 = 0,$$

hence $X^n(t)$ converges in distribution to $X^0(t)$, but since the distribution of $X^n(t)$ is the same as that of $X(t+\gamma_n)$ we deduce that $X(t+\gamma_n)$ converges in distribution to $X^0(t)$, i.e.

$$\lim_{n\to\infty} \mu(t+\alpha_n+\beta_n) = \operatorname{law}\big(X^0(t)\big) =: \mu_t^0.$$

By analogy and using (6), (7) we can easily deduce that

$$\lim_{n\to\infty}\lim_{m\to\infty} \mu(t+\alpha_n+\beta_m) = \mu_t^0.$$

We have thus proved that X has almost periodic one-dimensional distributions. To prove that X is almost periodic in distribution, we apply Proposition 3.2: for fixed $\tau\in\mathbb{R}$, let $\xi_n = X(\tau+\alpha_n)$, $F_n(t,x) = F(t+\alpha_n, x)$, $G_n(t,x) = G(t+\alpha_n, x)$. By the foregoing, (ξ_n) converges in distribution to some variable $Y(\tau)$. We can choose $Y(\tau)$ such that (ξ_n, W) converges in distribution to (Y, W). Then, for every $T\ge\tau$, $X(.+\alpha_n)$ converges in distribution on $C([\tau,T];\mathbb{H}_2)$ to the (unique in distribution) solution to

$$Y(t) = S(t-\tau)Y(\tau) + \int_\tau^t S(t-s)F\big(s, Y(s)\big)ds + \int_\tau^t S(t-s)G\big((s), Y(s)\big)dW(s).$$

Note that Y does not depend on the chosen interval $[\tau, T]$, thus the convergence takes place on $C(\mathbb{R};\mathbb{H}_2)$. Similarly, $Y_n := Y(.+\beta_n)$ converges in distribution on $C(\mathbb{R};\mathbb{H}_2)$ to a continuous process Z such that, for $t\ge\tau$,

$$Z(t) = S(t-\tau)Z(\tau) + \int_\tau^t S(t-s)F\big(s, Z(s)\big)ds + \int_\tau^t S(t-s)G\big((s), Z(s)\big)dW(s).$$

But, by (6) and (7), $X(.+\gamma_n)$ converges in distribution to the same process Z. Thus X is almost periodic in distribution. $\square$

4 Weak averaging

In this section we assume Conditions (i) – (iv) of Section 3, but we replace Condition (v) by Condition (v') below, which is weaker thanks to Propositions 2.4 and 2.5. Let us define the Hilbert space $\mathbb{H}_0 = \mathrm{range}(Q^{1/2})$ as in Proposition 2.5, where Q is the covariance operator of the Wiener process W. We assume that the mappings $F : \mathbb{R}\times\mathbb{H}_2 \to \mathbb{H}_2$ and $G : \mathbb{R}\times\mathbb{H}_2 \to L(\mathbb{H}_1, \mathbb{H}_2)$ satisfy

(v') There exists continuous functions $F_0 : \mathbb{H}_2 \to \mathbb{H}_2$ and $G_0 : \mathbb{H}_2 \to L(\mathbb{H}_0, \mathbb{H}_2)$ satisfying (2) and (3) for every $x \in \mathbb{H}_2$.

Theorem 4.1 *Let the assumptions* (i) – (iv) *and (v') be fulfilled and the constant* $\theta = \dfrac{K^2}{\delta^2}\Big(\dfrac{1}{2} + \mathrm{tr}\, Q\Big) < 1$. *For each fixed* $\varepsilon \in]0, 1[$, *let* X^ε *be the mild solution of the equation*

$$(10) \qquad dX^\varepsilon(t) = AX^\varepsilon(t)dt + F(\frac{t}{\varepsilon}, X^\varepsilon(t))dt + G(\frac{t}{\varepsilon}, X^\varepsilon(t))dW(t),\ t \in \mathbb{R}.$$

Then $(X^\varepsilon(t)) \to (X^0(t))$ *in distribution as* $\varepsilon \to 0+$ *on the space* $\mathcal{C}(\mathbb{R}, \mathbb{H}_2)$ *endowed with the topology of uniform convergence on compacts subsets of* $\mathbb{R}$, *where* $(X^0(t))$ *is the mild solution to*

$$(11) \qquad dX^0(t) = A(X^0(t))dt + F_0(X^0(t))dt + G_0(X^0(t))dW(t)$$

which is a stationary process.

Before we give the proof of this theorem, let us recall some well-known results.

Proposition 4.2 *([6]) Let* $(X_n)_{n\geq 0}$ *be a sequence of centered Gaussian random variable on a separable Hilbert space* $\mathbb{H}$ *with sequence of covariance operators* $(Q_n)_{n\geq 0}$. *Then* $(X_n)_{n\geq 0}$ *converges in distribution to* X_0 *in* $\mathbb{H}$ *if and only if*

$$\mid Q_n - Q_0 \mid_{\mathcal{N}} \to 0, n \to \infty$$

Let $\mathbb{U}, \mathbb{V}, \mathbb{H}$ be real separable Hilbert spaces, let W be a $\mathbb{U}$-valued $(\mathcal{F}_t)$-adapted Wiener process with nuclear covariance operator Q.

Proposition 4.3 *([17, Proposition 2.2]) Let $\alpha : \mathbb{H} \to \mathbb{V}$ be a Lipschitz mapping and $\sigma : \mathbb{R} \times \mathbb{H} \to L(\mathbb{U}, \mathbb{V})$ be a measurable mapping such that $\|\sigma(r, x)\|_{L(\mathbb{U},\mathbb{V})} \leq M(1 + \|x\|_{\mathbb{H}})$ and $\|\sigma(r, x) - \sigma(r, y)\|_{L(\mathbb{U},\mathbb{V})} \leq M\|x - y\|_{\mathbb{H}}$ for a constant M and every $r \in [s, t], x, y \in \mathbb{H}$. Let $g \in BL(\mathbb{V})$, we define*

$$\psi(y) = Eg\Big(\alpha(y) + \int_s^t \sigma(r, y)dW(r)\Big), y \in \mathbb{H}.$$

Let $u : \Omega \to \mathbb{H}$ be a $(\mathcal{F}_s)$-measurable random variable with $E\|u\|^2_{\mathbb{H}} < \infty$. Then

$$E\Big[g\Big(\alpha(u) + \int_s^t \sigma(r, u)dW(r)\Big)|\mathcal{F}_s\Big] = \psi(u) \ \mathrm{P}\text{-}a.s.$$

Proof of Theorem 4.1 We denote $F_\varepsilon(s, x) := F\big(\frac{s}{\varepsilon}, x\big)$, $G_\varepsilon(s, x) := G\big(\frac{s}{\varepsilon}, x\big)$, and for every $X \in \mathrm{CUB}\big(\mathbb{R}, L^2(\mathrm{P}, \mathbb{H}_2)\big)$,

$$L_\varepsilon(X)(t) := \int_{-\infty}^t S(t-s)F_\varepsilon(s, X(s))ds + \int_{-\infty}^t S(t-s)G_\varepsilon(s, X(s))dW(s),$$
$$L_0(X)(t) := \int_{-\infty}^t S(t-s)F_0(X(s))ds + \int_{-\infty}^t S(t-s)G_0(X(s))dW(s).$$

First step. Let us show that $L_\varepsilon(X) \longrightarrow L_0(X)$ in distribution, as $\varepsilon \to 0$, in the space $\mathcal{C}(\mathbb{R}, \mathbb{H}_2)$ endowed with the topology of uniform convergence on the compact subsets of $\mathbb{R}$. This amounts to prove that, for any $\tau \in \mathbb{R}$ and any $T \in \mathbb{R}$ such that $\tau \leq T$, $L_\varepsilon(X) \longrightarrow L_0(X)$ in distribution in the space $\mathcal{C}([\tau, T], \mathbb{H}_2)$ (see [18, Theorem 5]). Since the previous integral exists, then for every $\eta > 0$, there exists τ such that for each $\sigma < \tau$, we have $E\|Y^n(\sigma)\|^2_{\mathbb{H}_2} < \eta$, thus for the proof that Y^n converges in distribution to Y^0 on $\mathcal{C}(\mathbb{R}, H_2)$, it suffices to show the convergence in distribution on $\mathcal{C}([\tau, T], \mathbb{H}_2)$ of

$$\int_{-\tau}^t S(t-s)F_\varepsilon(s, X(s))ds + \int_{-\tau}^t S(t-s)G_\varepsilon(s, X(s))dW(s)$$

to

$$\int_{-\tau}^t S(t-s)F_0(X(s))ds + \int_{-\tau}^t S(t-s)G_0(X(s))dW(s).$$

Let (ε_n) be an arbitrary sequence in $]0, 1[$ such that $\varepsilon_n \longrightarrow 0$ and $X \in \mathrm{CUB}\big(\mathbb{R}, L^2(\mathrm{P}, \mathbb{H}_2)\big)$. We denote $F_n(s, x) := F(\frac{s}{\varepsilon_n}, x)$, $G_n(s, x) := G(\frac{s}{\varepsilon_n}, x)$,

we simplify the notation L_{ε_n} in L_n, and we denote

$$\begin{aligned}Y^n(t) &:= L_n(X)(t)\\ &:= \int_{-\infty}^t S(t-s)F_n(s,X(s))ds + \int_{-\infty}^t S(t-s)G_n(s,X(s))dW(s).\end{aligned}$$

We have $\sup E\|X_t\|^2_{\mathbb{H}_2} < \infty$, thus the process X satisfies the following condition: for every $\eta > 0$, there exist a partition

$$\{\tau = t_o < t_1 < \cdots < t_k = T\} \text{ of } [\tau, T]$$

and a process

$$\widetilde{X}(t) = \Sigma_{i=1}^k \widetilde{X}(t_{i-1}) 1_{[t_{i-1},t_i[}(t)$$

such that

$$\sup E\|X(t) - \widetilde{X}(t)\|^2_{\mathbb{H}_2} < \eta.$$

Using the fact that L_ε is Lipschitz, we obtain

$$\sup E\|L_\varepsilon X(t) - L_\varepsilon \widetilde{X}(t)\|^2_{\mathbb{H}_2} < \eta$$

uniformly with respect to ε.

We set, with a slight abuse of notation:

$$\begin{aligned}\widetilde{X}^n(t) = L_n \widetilde{X}(t) = \sum_{i=1}^k \Big(&\int_{t_{i-1}\wedge t}^{t_i \wedge t} S(t-s)F_n(s, \widetilde{X}_{t_{i-1}})ds\\ &+ \int_{t_{i-1}\wedge t}^{t_i\wedge t} S(t-s)G_n(s,\widetilde{X}_{t_{i-1}})dW(s)\Big)\end{aligned}$$

To prove that for each $t \in [\tau, T]$, $L_n X(t)$ converges in distribution to $L_0 X(t)$, it suffices to show that for each $l \in \{0, 1, \dots, k\}$, $\widetilde{X}^n(t_l) \longrightarrow \widetilde{X}^0(t_l)$ in distribution as $n \to \infty$, where

$$\begin{aligned}\widetilde{X}^n(t_l) =& \sum_{i=1}^l \Big(\int_{t_{i-1}}^{t_i} S(t_l - s)F_n(s, \widetilde{X}(t_{i-1}))ds\\ &+ \int_{t_{i-1}}^{t_i} S(t_l - s)G_n(s, \widetilde{X}(t_{i-1}))dW(s)\Big)\\ =& S(t_l - t_{l-1})\widetilde{X}^n(t_{l-1}) + \int_{t_{l-1}}^{t_l} S(t_l - s)F_n(s, \widetilde{X}(t_{l-1}))ds\\ &+ \int_{t_{l-1}}^{t_l} S(t_l - s)G_n(s, \widetilde{X}(t_{l-1}))dW(s).\end{aligned}$$

Indeed, we just use the following two properties: For any $\gamma > 0$ there exists a partition $\{\tau = t_0 < \cdots < t_k = T\}$ of the interval $[\tau, T]$ such that

$$\sup_n E\Big(\max_{l=1,\dots,k+1} \max_{t_{l-1} \le t \le t_l} \Big| \widetilde{X}^n(t) - \widetilde{X}^n(t_{l-1}) \Big| \Big) < \gamma$$

and

$$\begin{aligned} d_{\mathrm{BL}}\Big(\mathrm{law}\,(L_n X(t))\,,\mathrm{law}\,(L_0 X(t))\Big) \le & d_{\mathrm{BL}}\Big(\mathrm{law}\,(L_n X(t))\,,\mathrm{law}\,\Big(\widetilde{X}^n(t)\Big)\Big) \\ & + d_{\mathrm{BL}}\Big(\mathrm{law}\,\Big(\widetilde{X}^n(t)\Big)\,,\mathrm{law}\,\Big(\widetilde{X}^0(t)\Big)\Big) \\ & + d_{\mathrm{BL}}\Big(\mathrm{law}\,\Big(\widetilde{X}^0(t)\Big)\,,\mathrm{law}\,(L_0 X(t))\Big) \\ \le & E\|L_n X(t) - \widetilde{X}^n(t)\| \\ & + d_{\mathrm{BL}}\Big(\mathrm{law}\,\Big(\widetilde{X}^n(t)\Big)\,,\mathrm{law}\,\Big(\widetilde{X}^0(t)\Big)\Big) \\ & + E\|L_0 X(t) - \widetilde{X}^0(t)\|. \end{aligned}$$

We define a mapping

$$\gamma_n : \mathbb{H}_2^l \longrightarrow L^1(\Omega, \mathbb{H}_2)$$

by

$$\begin{aligned} \gamma_n(y_0, y_1, \dots, y_{l-1}) = \sum_{i=1}^{l} \Big(& \int_{t_{i-1}}^{t_i} S(t_l - s) F_n(s, y_{i-1})\, ds \\ & + \int_{t_{i-1}}^{t_i} S(t_l - s) G_n(s, y_{i-1})\, dW(s) \Big) \end{aligned}$$

Obviously,

$$\gamma_n(\widetilde{X}_{t_0}, \widetilde{X}_{t_1}, \dots, \widetilde{X}_{t_{l-1}}) = \widetilde{X}^n_{t_l}$$

Let

$$\mu_{t_0, t_1, \dots, t_{l-1}} = \mathrm{law}\,\Big(\widetilde{X}_{t_0}, \widetilde{X}_{t_1}, \dots, \widetilde{X}_{t_{l-1}}\Big)\,.$$

Let $g \in BL(\mathbb{H}_2)$, and $h_n(y) = E[g(\gamma_n(y))]$; $y \in \mathbb{H}_2^l$. From Proposition 4.3, we have

$$E[g(\widetilde{X}^n_{t_l})] = E[h_n(\widetilde{X}_{t_0}, \widetilde{X}_{t_1}, \dots, \widetilde{X}_{t_{l-1}})] = \int_{\mathbb{H}_2^l} h_n(y) d\mu_{t_0, t_1, \dots, t_{l-1}}(y)$$

and

$$|E[g(\widetilde{X}^n_{t_l})]-E[g(\widetilde{X}^0_{t_l})]| = |\int_{\mathbb{H}_2^l} h_n(y)-h_0(y)d\mu_{t_0,t_1,\dots,t_{l-1}}(y)|$$
$$\leq \int_{\mathbb{H}_2^l} |h_n(y)-h_0(y)|d\mu_{t_0,t_1,\dots,t_{l-1}}(y).$$

Let us show that $h_n(y) \longrightarrow h_0(y)$ as $n \to \infty$ for every $y \in \mathbb{H}_2^l$:

$$\begin{aligned}\gamma_n(y)-\gamma_0(y) =&\sum_{i=1}^{l}\int_{t_{i-1}}^{t_i} S(t_l-s)\big(F_n(s,y_{i-1})-F_0(y_{i-1})\big)ds\\ &+\sum_{i=1}^{l}\int_{t_{i-1}}^{t_i} S(t_l-s)\big(G_n(s,y_{i-1})-G_0(y_{i-1})\big)\,dW(s)\\ =&I_n+J_n.\end{aligned}$$

Assumption (v') implies that $I_n \longrightarrow 0$ as $n \longrightarrow \infty$, and since

$$\sum_{i=1}^{l}\int_{t_{i-1}}^{t_i} S(t_l-s)G_n(s,y_{i-1})\,dW(s)$$

is a centered Gaussian random variable in $\mathbb{H}_2$, we deduce by Assumption (v') and Proposition 4.2 that $J_n \longrightarrow 0$ in distribution as $n \longrightarrow \infty$ hence $\gamma_n(y) \longrightarrow \gamma_0(y)$ in distribution as $n \longrightarrow \infty$, consequently

$$h_n(y) \longrightarrow h_0(y) \text{ for any } y \in \mathbb{H}_2^l. \tag{12}$$

For every $\eta' > 0$, there exists a compact set $\mathcal{K} \subset \mathbb{H}_2^l$ such that

$$\mu_{t_0,t_1,\dots,t_{l-1}}(\mathbb{H}_2^l \setminus \mathcal{K}) < \eta'.$$

We have

$$h_n \in BL(\mathbb{H}_2^l) \text{ and } \sup_n \|h_n\|_{\mathrm{BL}} < \infty \tag{13}$$

because, for all $y,z \in \mathbb{H}_2^l$, and for some constant K_1,

$$|h_n(y)-h_n(z)| \leq \|g\|_{\mathrm{BL}} E\|\gamma_n(y)-\gamma_n(z)\|_{\mathbb{H}_2} \leq K_1\|g\|_{\mathrm{BL}}\|y-z\|_{\mathbb{H}_2^l}.$$

From (12), (13) and the compactness of $\mathcal{K}$, we deduce that h_n converges to h_0 uniformly on $\mathcal{K}$, hence

$$\lim_{n\to\infty}\int_{\mathcal{K}} |h_n(y)-h_0(y)|d\mu_{t_0,t_1,\dots,t_{l-1}}(y) = 0$$

and, since g is a bounded function,

$$\int_{\mathbb{H}_2^l \setminus \mathcal{K}} |h_n(y) - h_0(y)| d\mu_{t_0,t_1,\ldots,t_{l-1}}(y)$$
$$\leq 2 \sup_n \sup_y |h_n(y)| \eta' = 2 \sup_n \sup_y |E[g(\gamma_n(y))]| \eta'.$$

Thus $\widetilde{X}^n(t_l) \longrightarrow \widetilde{X}^0(t_l)$ in distribution as $n \to \infty$.

Let us now show that $(\widetilde{X}^n_{t_0}, \widetilde{X}^n_{t_1}, \ldots, \widetilde{X}^n_{t_k}) \longrightarrow (\widetilde{X}^0_{t_0}, \widetilde{X}^0_{t_1}, \ldots, \widetilde{X}^0_{t_k})$ in distribution as $n \longrightarrow \infty$. We proceed by induction. By the foregoing, we have $\widetilde{X}^n_{t_0} \longrightarrow \widetilde{X}^0_{t_0}$ in distribution. Assume that for $0 \leq l \leq k-1$, $(\widetilde{X}^n_{t_0}, \widetilde{X}^n_{t_1}, \ldots, \widetilde{X}^n_{t_l})$ converges in distribution in $\mathbb{H}_2^{l+1}$. Let us define $\alpha_n : \mathbb{H}_2^{l+1} \longrightarrow \mathbb{H}_2^{l+2}$ by

$$\alpha_n(y_0, y_1, \ldots, y_l) = \Big(y_0, y_1, \ldots, y_l, S(t_{l+1}-t_l)y_l + \int_{t_l}^{t_{l+1}} S(t_{l+1}-s) F_n(s, y_l) ds\Big)$$

and $\beta_n : \mathbb{H}_2^{l+1} \longrightarrow L^1(\Omega, \mathbb{H}_2^{l+2})$ by

$$\beta_n(y_0, y_1, \ldots, y_l) = \Big(0, \ldots, 0, \int_{t_l}^{t_{l+1}} S(t_{l+1}-s) G_n(s, y_l) dW(s)\Big)$$

so that

$$(\alpha_n + \beta_n)(\widetilde{X}^n_{t_0}, \widetilde{X}^n_{t_1}, \ldots, \widetilde{X}^n_{t_l}) = (\widetilde{X}^n_{t_0}, \widetilde{X}^n_{t_1}, \ldots, \widetilde{X}^n_{t_l}, \widetilde{X}^n_{t_{l+1}}).$$

We denote $u_n = (\widetilde{X}^n_{t_0}, \widetilde{X}^n_{t_1}, \ldots, \widetilde{X}^n_{t_l})$ and $\mu_n = \mathrm{law}\,(u_n)$. Let $g \in BL(\mathbb{H}_2^{l+2})$, and

$$h_n(y) = Eg\big(\alpha_n(y) + \beta_n(y)\big), \ y \in \mathbb{H}_2^{l+1}.$$

Proposition 4.3 yields

$$Eg(\widetilde{X}^n_{t_0}, \widetilde{X}^n_{t_1}, \ldots, \widetilde{X}^n_{t_{l+1}}) = Eh_n(u_n) = \int_{\mathbb{H}_2^{l+1}} h_n(y) d\mu_n(y).$$

It follows that

$$\begin{aligned}
&|Eg(\widetilde{X}^n_{t_0}, \widetilde{X}^n_{t_1}, \ldots, \widetilde{X}^n_{t_{l+1}}) - Eg(\widetilde{X}^0_{t_0}, \widetilde{X}^0_{t_1}, \ldots, \widetilde{X}^0_{t_{l+1}})| \\
&\leq \int_{\mathbb{H}_2^{l+1}} |h_n(y) - h_0(y)| d\mu_n(y) + \Big|\int_{\mathbb{H}_2^{l+1}} h_0(y) d\mu_n(y) - \int_{\mathbb{H}_2^{l+1}} h_0(y) d\mu_0(y)\Big| \\
&\leq J_1(n) + J_2(n).
\end{aligned}$$

As in the above reasoning, we can prove that

$$h_n \in BL(\mathbb{H}_2^{l+1}),\ \sup_n \|h_n\|_{\text{BL}} < \infty$$

and

$$\alpha_n(y) + \beta_n(y) \longrightarrow \alpha_0(y) + \beta_0(y)$$

in distribution, for every $y \in \mathbb{H}_2^{l+1}$ thus $h_n(y) \longrightarrow h_0(y)$ as $n \longrightarrow \infty$, for any $y \in \mathbb{H}_2^{l+1}$. On the other hand, by the induction hypothesis we have $\mu_n \longrightarrow \mu_0$ and since $h_0 \in BL(\mathbb{H}_2^{l+1})$ we have $J_2(n) \longrightarrow 0$ as $n \longrightarrow 0$. The convergence of μ_n implies that $\{\mu_n\}$ is tight, i.e. for each $\eta^{'} > 0$ there exists a compact set $\mathcal{K} \in \mathbb{H}_2^{l+1}$ such that

$$\sup_n \mu_n(\mathbb{H}_2^{l+1} \setminus \mathcal{K}) < \eta^{'} \tag{14}$$

Since for every $y \in \mathbb{H}_2^{l+1}$, $h_n(y) \longrightarrow h_0(y)$, from (14) and the compactness of $\mathcal{K}$ the function h_n converges to h_0 uniformly on $\mathcal{K}$, hence

$$\lim_{n\to\infty} \int_{\mathcal{K}} |h_n(y) - h_0(y)| d\mu_n(y) = 0$$

and

$$\int_{\mathbb{H}_2^l \setminus \mathcal{K}} |h_n(y) - h_0(y)| d\mu_n(y) \leq 2 \sup_n \sup_y |h_n(y)| \eta$$

So $J_1(n) \longrightarrow 0$ as $n \longrightarrow 0$, consequently $(\widetilde{X}_{t_0}^n, \widetilde{X}_{t_1}^n, \dots, \widetilde{X}_{t_{l+1}}^n)$ converges in distribution in $\mathbb{H}_2^{l+2}$, which implies that for every $q \in \mathbb{N}$, we have

$$(Y_{t_0}^n, Y_{t_1}^n, \dots, Y_{t_q}^n) \to (Y_{t_0}^0, Y_{t_1}^0, \dots, Y_{t_q}^0) \tag{15}$$

in distribution on $\mathbb{H}_2^{q+1}$.

It remains to show that for each positive η, ν there exist α, $0 < \alpha < 1$ and an integer n_0, such that

$$\frac{1}{\alpha} \mathrm{P}\{ \sup_{s \leq t \leq s+\alpha} \|Y^n(s) - Y^n(t)\| > \nu\} < \eta,\ n \geq n_0 \tag{16}$$

for every s in $[\tau, T]$. We have

$$\begin{aligned}
&Y^n(t) - Y^n(s) \\
&= \int_\tau^t S(t-\sigma) F_n(\sigma, X(\sigma)) d\sigma - \int_\tau^s S(s-\sigma) F_n(\sigma, X(\sigma)) d\sigma \\
&\quad + \int_\tau^t S(t-\sigma) G_n(\sigma, X(\sigma)) dW(\sigma) - \int_\tau^s S(s-\sigma) G_n(\sigma, X(\sigma)) dW(\sigma) \\
&= \int_s^t S(t-\sigma) F_n(\sigma, X(\sigma)) d\sigma + \int_s^t S(t-\sigma) G_n(\sigma, X(\sigma)) dW(\sigma).
\end{aligned}$$

We have

$$\frac{1}{\alpha}\mathrm{P}\{\sup_{s\le t\le s+\alpha}\|Y^n(s)-Y^n(t)\|>\nu\}\le\frac{1}{\alpha\nu^4}E(\sup_{s\le t\le s+\alpha}\|Y^n(s)-Y^n(t)\|^4)$$

To check Inequality (16) it suffices to show that we can choose α such that

$$E(\sup_{s\le t\le s+\alpha}\|Y^n(s)-Y^n(t)\|^4)\le\alpha\nu^4\eta$$

for every integer n and every $s\in[\tau,T]$.

Using the inequality $(a+b)^4\le 8a^4+8b^4$, we obtain

$$\begin{aligned}E(\sup_{s\le t\le s+\alpha}\|Y^n(s)-Y^n(t)\|^4)\le&8E(\sup_{s\le t\le s+\alpha}\|\int_s^t S(t-\sigma)F_n(\sigma,X(\sigma))d\sigma\|^4)\\&+8E(\sup_{s\le t\le s+\alpha}\|\int_s^t S(t-\sigma)G_n(\sigma,X(\sigma))dW(\sigma)\|^4)\\\le&I_1+I_2.\end{aligned}$$

Applying Cauchy-Schwarz inequality, we get

$$\begin{aligned}I_1&=8E(\sup_{s\le t\le s+\alpha}\|\int_s^t S(t-\sigma)F_n(\sigma,X(\sigma))d\sigma\|^4)\\&\le 8E\Big(\sup_{s\le t\le s+\alpha}\Big(\int_s^t\|S(t-\sigma)F_n(\sigma,X(\sigma)\|d\sigma\Big)^4\Big)\\&\le 8E\sup_{s\le t\le s+\alpha}\Big[\Big(\int_s^t e^{-2\delta(t-\sigma)}d\sigma\Big)^2\Big(\int_s^t\|F_n(\sigma,X(\sigma)\|^2d\sigma\Big)^2\Big]\\&\le 8K^2\sup_{s\le t\le s+\alpha}\Big(\frac{1}{2\delta}(1-e^{-2\delta(t-s)})\Big)^2E\Big(\int_s^{s+\alpha}(1+\|X(\sigma)\|)^2d\sigma\Big)^2\end{aligned}$$

For I_2, using the stochastic convolution inequality, we obtain, for some constant C_{conv},

$$\begin{aligned}I_2&=8E(\sup_{s\le t\le s+\alpha}\|\int_s^t S(t-\sigma)G_n(\sigma,X(\sigma))dW(\sigma)\|^4)\\&\le 8C_{\mathrm{conv}}E\Big(\int_s^{s+\alpha}\|G_n(\sigma,X(\sigma))\|^2d\sigma\Big)^2\\&\le 8C_{\mathrm{conv}}K^2\int_s^{s+\alpha}(1+\|X(\sigma)\|)^2d\sigma.\end{aligned}$$

Thus, for α small enough,

$$E\big(\sup_{s\le t\le s+\alpha}\|Y^n(s)-Y^n(t)\|^2\big)\le\eta.$$

Therefore $L_\varepsilon(X)\to L_0(X)$ in distribution as $\varepsilon\to 0+$ on the space $\mathcal{C}([\tau,T],\mathbb{H}_2)$ for all τ,T such that $T>\tau$. Hence $L_\varepsilon(X)\to L_0(X)$ in distribution as $\varepsilon\to 0+$ on the space $\mathcal{C}(\mathbb{R},\mathbb{H}_2)$.

Second step. Now, let us show that $L_\varepsilon(X^\varepsilon)\to L_0(X^0)$ in distribution as $\varepsilon\to 0+$ on the space $\mathcal{C}(\mathbb{R},\mathbb{H}_2)$, which means that $X^\varepsilon\to X^0$ in distribution as $\varepsilon\to 0+$ on the space $\mathcal{C}(\mathbb{R},\mathbb{H}_2)$.

We denote by: $\mu_\varepsilon^k(X):=\operatorname{law}\big(L_\varepsilon^k(X)\big)$ $(k\ge 1)$ where $L_\varepsilon^k=L_\varepsilon\circ L_\varepsilon\circ\cdots\circ L_\varepsilon$, $\mu(X):=\operatorname{law}(X)$ and $\mu_0^k(X):=\operatorname{law}\big(L_0^k(X)\big)$ where $L_0^k=L_0\circ L_0\circ\cdots\circ L_0$.

Observe that:

(a) Since, for every $k\in\mathbb{N}$, L_ε^k and L_0^k are contraction operators (see the proof of Theorem 3.1), we deduce that, for every $f\in BL(\mathcal{C}(\mathbb{R},\mathbb{H}_2)$ such that $\|f\|_{\rm BL}\le 1$, for each $\omega\in\Omega$, $f\circ L_\varepsilon^k(\omega)\in BL(\mathcal{C}(\mathbb{R},\mathbb{H}_2)$ and $\|f\circ L_\varepsilon^k(w)\|_{\rm BL}\le 1$.

(b) Since X^ε and X^0 are solutions, we have $\mu_\varepsilon^k(X^\varepsilon)=\mu(X^\varepsilon)$ and $\mu_0^k(X^0)=\mu(X^0)$ for every $k\in\mathbb{N}^*$

(c) Since L_ε is θ-Lipschitz,

$$\Big(E\|L_\varepsilon^k(X)-L_\varepsilon^k(Y)\|\Big)^2\le E\|L_\varepsilon^k(X)-L_\varepsilon^k(Y)\|^2\le\theta^kE\|X-Y\|^2$$

Now, using the properties of the metric $d_{\rm BL}$ on the probability space $\mathcal{P}\left(\mathcal{C}(\mathbb{R},\mathbb{H}_2)\right)$ and (c) we obtain, for $X,Y\in\mathrm{CUB}\big(\mathbb{R},L^2(\mathrm{P},\mathbb{H}_1)\big)$,

$$\begin{aligned}d_{\rm BL}(\mu_\varepsilon^k(X),\mu_\varepsilon^k(Y))&=\sup_{\|f\|_{\rm BL}\le1}\Big|\int_{\mathcal{C}(\mathbb{R},\mathbb{H}_2)}fd(\mu_\varepsilon^k(X)-\mu_\varepsilon^k(Y))\Big|\\&=\sup_{\|f\|_{\rm BL}\le1}\Big|\int_\Omega f[L_\varepsilon^k(X)]-f[L_\varepsilon^k(Y)]d\,\mathrm{P}\Big|\\&\le\int_\Omega\|L_\varepsilon^k(X)-L_\varepsilon^k(Y)\|d\,\mathrm{P}\le\theta^kE\|X-Y\|^2\end{aligned}$$

Furthermore, from (a) we obtain

$$\begin{aligned}d_{\rm BL}(\mu_\varepsilon^k(X),\mu_\varepsilon^k(Y))&=\sup_{\|f\|_{\rm BL}\le1}\Big|\int_\Omega f[L_\varepsilon^k(X)]-f[L_\varepsilon^k(Y)]d\,\mathrm{P}\Big|\\&\le\sup_{\|f\|_{\rm BL}\le1}\Big|\int_\Omega f(X)-f(Y)d\,\mathrm{P}\Big|=d_{\rm BL}\big(\mu(X),\mu(Y)\big)\end{aligned}\tag{17}$$

From above we deduce that there exist $k \in \mathbb{N}^*$ and $0 < \theta^{'} < 1$ such that

$$d_{\mathrm{BL}}(\mu_\varepsilon^k(X), \mu_\varepsilon^k(Y)) \le \theta^{'} d_{\mathrm{BL}}\big(\mu(X), \mu(Y)\big) \tag{18}$$

Indeed, assume that Inequality (18) is false, for every $0 < \theta^{'} < 1$ there exist $X, Y \in \mathrm{CUB}\big(\mathbb{R}, L^2(\mathrm{P}, \mathbb{H}_1)\big)$ such that $\mu(X) \ne \mu(Y)$ and

$$\theta^k \Big(E\|X - Y\|^2\Big)^{\frac{1}{2}} \ge d_{\mathrm{BL}}(\mu_\varepsilon^k(X), \mu_\varepsilon^k(Y)) > \theta^{'} d_{\mathrm{BL}}\big(\mu(X), \mu(Y)\big);\ \forall k \in \mathbb{N}^*$$

therefore taking k elarge enough, we see that $d_{\mathrm{BL}}\big(\mu(X), \mu(Y)\big) = 0$, a contradiction with $\mu(X) \ne \mu(Y)$.

Now, by using (b), (18), and (17), it is easy to conclude that

$$\begin{aligned} d_{\mathrm{BL}}\big(\mu(X^\varepsilon), \mu(X^0)\big) &= d_{\mathrm{BL}}(\mu_\varepsilon^k(X^\varepsilon), \mu_0^k(X^0) \\ &\le \frac{1}{1-\theta^{'}} d_{\mathrm{BL}}(\mu_\varepsilon^k(X^0), \mu_0^k(X^0) \\ &\le \frac{k-1}{1-\theta^{'}} d_{\mathrm{BL}}(\mu_\varepsilon^1(X^0), \mu_0^1(X^0). \end{aligned}$$

Indeed, we have

$$\begin{aligned} (1-\theta^{'}) & d_{\mathrm{BL}}\big(\mu(X^\varepsilon), \mu(X^0)\big) \\ &= d_{\mathrm{BL}}\big(\mu(X^\varepsilon), \mu(X^0)\big) - \theta^{'} d_{\mathrm{BL}}\big(\mu(X^\varepsilon), \mu(X^0)\big) \\ &\le d_{\mathrm{BL}}\big(\mu(X^\varepsilon), \mu(X^0)\big) - d_{\mathrm{BL}}(\mu_\varepsilon^k(X^\varepsilon), \mu_\varepsilon^k(X^0)) \\ &= d_{\mathrm{BL}}\big(\mu_\varepsilon^k(X^\varepsilon), \mu(X^0)\big) - d_{\mathrm{BL}}(\mu_\varepsilon^k(X^\varepsilon), \mu_\varepsilon^k(X^0)) \\ &\le d_{\mathrm{BL}}(\mu_\varepsilon^k(X^0), \mu(X^0)) = d_{\mathrm{BL}}(\mu_\varepsilon^k(X^0), \mu_0^1(X^0)) \\ &\le d_{\mathrm{BL}}(\mu_\varepsilon^k(X^0), \mu_\varepsilon^{k-1}(X^0)) + d_{\mathrm{BL}}(\mu_\varepsilon^{k-1}(X^0), \mu_0^1(X^0)). \end{aligned} \tag{19}$$

We thus have

$$\begin{aligned} (1-\theta^{'}) & d_{\mathrm{BL}}\big(\mu(X^\varepsilon), \mu(X^0)\big) \\ &\le d_{\mathrm{BL}}(\mu_\varepsilon^{k-1}(L_\varepsilon(X^0)), \mu_\varepsilon^{k-1}(X^0)) + d_{\mathrm{BL}}(\mu_\varepsilon^{k-1}(X^0), \mu_0^1(X^0)) \\ &\le d_{\mathrm{BL}}(\mu_\varepsilon^1(X^0), \mu(X^0)) + d_{\mathrm{BL}}(\mu_\varepsilon^{k-1}(X^0), \mu_0^1(X^0)) \\ &= d_{\mathrm{BL}}(\mu_\varepsilon^1(X^0), \mu_0^1(X^0)) + d_{\mathrm{BL}}(\mu_\varepsilon^{k-1}(X^0), \mu_0^1(X^0)) \\ &\le (k-1) d_{\mathrm{BL}}(\mu_\varepsilon^1(X^0), \mu_0^1(X^0)) \end{aligned}$$

the last inequality being obtained by finite induction, repeating the calculation of (19). Thus, using the result of the first step,

$$\lim_{\varepsilon \to 0} d_{\mathrm{BL}}\big(\mu(X^\varepsilon), \mu(X^0)\big) = 0.$$

Finally, by [8, Theorem 4.1] we deduce that the mild solution (X^0) of (11) is a stationary process. $\square$

References

[1] Paul H. Bezandry and Toka Diagana. Existence of almost periodic solutions to some stochastic differential equations. *Appl. Anal.*, 86(7):819–827, 2007.

[2] Paul H. Bezandry and Toka Diagana. Square-mean almost periodic solutions nonautonomous stochastic differential equations. *Electron. J. Differential Equations*, pages No. 117, 10 pp. (electronic), 2007.

[3] Paul H. Bezandry and Toka Diagana. Existence of quadratic-mean almost periodic solutions to some stochastic hyperbolic differential equations. *Electron. J. Differential Equations*, pages No. 111, 14, 2009.

[4] N. N. Bogoliubov and Y. A. Mitropolsky. *Asymptotic methods in the theory of non-linear oscillations.* Translated from the second revised Russian edition. International Monographs on Advanced Mathematics and Physics. Hindustan Publishing Corp., Delhi, Gordon and Breach Science Publishers, New York, 1961.

[5] V. Sh. Burd. *Method of Averaging for Differential Equations on an Infinite Interval: Theory and Applications*, volume 255 of *Lecture Notes in Pure and Applied Mathematics.* Chapman and Hall/CRC; 1 edition, 2007.

[6] Simone Chevet. Compacité dans l'espace des probabilités de Radon gaussiennes sur un Banach. *C. R. Acad. Sci. Paris Sér. I Math.*, 296(5):275–278, 1983.

[7] Constantin Corduneanu. *Almost periodic oscillations and waves.* Springer, New York, 2009.

[8] G. Da Prato and C. Tudor. Periodic and almost periodic solutions for semilinear stochastic equations. *Stochastic Anal. Appl.*, 13(1):13–33, 1995.

[9] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions.* Number 44 in Encyclopedia of Mathematics and Its Applications. Cambridge University Press, Cambridge, 1992.

[10] A. M. Fink. *Almost periodic differential equations.* Lecture Notes in Mathematics, Vol. 377. Springer-Verlag, Berlin, 1974.

[11] R. Z. Khasminskii. On the principle of averaging the itô's stochastic differential equations. *Kybernetika (Prague)*, 4:260–279, 1968.

[12] M. A. Krasnosel′skiĭ, V. Sh. Burd, and Yu. S. Kolesov. *Nonlinear almost periodic oscillations.* Halsted Press [A division of John Wiley & Sons], New York-Toronto, Ont., 1973. Translated from the Russian by A. Libin, Translation edited by D. Louvish.

[13] N. Krylov and N. Bogoliubov. *Introduction to non-linear mechanics.* Annals of Mathematics Studies. 11. Princeton, N.J.: Princeton University Press, III, 105 p. , 1943.

[14] O. Mellah and P. Raynaud de Fitte. Counterexamples to mean square almost periodicity of the solutions of some SDEs with almost periodic coefficients. *ArXiv e-prints*, August 2012, 1208.6384.

[15] Claudia Prévôt and Michael Röckner. *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics.* Springer, Berlin, 2007.

[16] C. Tudor. Almost periodic stochastic processes. In *Qualitative problems for differential equations and control theory*, pages 289–300. World Sci. Publ., River Edge, NJ, 1995.

[17] Ivo Vrkoč. Weak averaging of stochastic evolution equations. *Math. Bohem.*, 120(1):91–111, 1995.

[18] Ward Whitt. Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Math. Statist.*, 41:939–944, 1970.